基于多变量LSTM网络的太阳黑子活动预测分析[1]

梁波[1,2]，林语琦[1,2]，戴伟[1,2]，冯松[1,2]，杨云飞[1,2]

(1.昆明理工大学信息工程与自动化学院，云南 昆明 650500；

2.云南省计算机应用重点实验室，云南 昆明 650500)

摘要：通过添加太阳黑子周期长度构建长短期记忆网络（Long Short-Term Memory，LSTM)的多变量输入数据，在多时间步长上预测未来10年的太阳黑子变化。将数据集以训练数据长度为标准划分出两组时间序列片段，分别是分片11和分片6，并在分片上分别对比了单变量和多变量在单时间步长和多时间步长上的预测效果。最后，得出以下主要结论：(1)比起分片6，分片11的采样方式有更低的均方根误差(Root mean squared error,RMSE)；(2)最优多步长的预测效果比单步长的要好；(3)通过图例上周期的起始点、结束点和最大振幅位置三个点的准确度证明了多变量的多步长方法确实有更好的预测效果。

关键字：太阳黑子；长短期记忆网络；多变量；时间步长；均方根误差

中图分类号：TP391.41　　**文献标识码：**A　　**文章编号：**1672-7673（2020）03

太阳黑子是日面上最常见的一种太阳活动现象，它们的数量随时间的变化呈现大约11年的周期变化。太阳黑子数的增加往往伴随着太阳爆发活动的增强。灾害性空间天气则会引起地球磁场的扰动，干扰雷达和无线电通讯，同时也间接影响地球的气候变化[1-2]。基于已　是预测下一个太阳活动周的活动指数的最大振幅、持续的时间以及最大振幅出现的时间。

太阳活动的预测方法主要有3类：前体预测法、外推方法和模型预测法。在前体预测方法中，无论是采用地磁度量值[3]预测还是极地测量值[4]预测太阳活动的最大幅度，预测结果都普遍存在延迟。外推方法包括线性回归和非线性回归法[5-8]。模型预报法中的模型主要是指太阳发电机组[9-10]，是建立在大量物理基础之上的物理模型，详见文[11]。

目前，第24个太阳活动周已经接近尾声，所以太阳活动周的预测都是围绕第25周展开[12-15]。随着机器学习和深度学习技术的出现，许多方法被应用于25活动周的预测分析中[16-18]。其中，文[16]采用长短期记忆网络模型预测了第25个太阳活动周期的峰值出现在2022年7月，介绍了如何在一个很长的太阳黑子序列中开发预测模型，并将预测结果与真实观测值之间的均方根误差作为模型预测效果的参考依据。

长短期记忆网络最早是由文[19]提出，后发展为循环神经网络(Recurrent Neural Network，RNN)中的一种。长短期记忆网络的工作单元由“门”保持或抑制单元状态，以及单元内外的信息交流，单元之间按链式链接的特殊结构为数据之间提供长期的依赖关系，这种信息持久化的特性被广泛应用到时间序列的相关工作中。长短期记忆网络的递归特性使得单元不会立即输出当前输入的响应，而是等待时间间隙的到来，时间间隙由时间步长控制。文[16]采用长短期记忆网络模型和时间步长为1的单变量预测方法。

本文提出了一种构建多变量的方法，结合长短期记忆网络预测未来10年的

1基金项目：国家自然科学联合基金（U1931107）；云南省应用基础研究计划重点项目（2018FA035）资助.

收稿日期：2019-11-27；修订日期：2020-01-17

作者简介：梁 波，男，工程师.研究方向：天文数据处理. Email：liangbo@astrolab.cn;

通讯作者：冯 松，男，教授.研究方向：天文数据处理. Email: feng.song@astrolab.cn

太阳黑子数，并对比了单变量和多变量分别在单时间步长和多种多时间步长的预测结果，最后得出最优策略预测未来 10 年的太阳黑子数，该策略由预测结果的最小均方根误差得出。结果证明多变量在多步长预测上的效果更好，并且用图例上 3 个点的准确度证明了本文的结论，3 个点分别是周期的起始点、结束点和最大振幅位置。

1. 数据

太阳黑子数据采用 SILSO 网站(http://www.sidc.be/silso/datafiles)发布的太阳黑子 2.0 版本中的 13 个月平滑月度数据(Source: WDC-SILSO, Royal Observatory of Belgium, Brussels)。数据范围从 1749 年 6 月到 2019 年 3 月，历时 270 年，包含 3237 组数据，如图 1，图中的横坐标为每个太阳周期的起始日期，由太阳黑子周期记录(http://sidc.oma.be/silso/cyclesmm)给出。例如第 1 个周期的时间范围是 1755 年 2 月至 1766 年 5 月，第 2 个周期是 1766 年 6 月到 1775 年 5 月。下文将 13 个月平滑月度太阳黑子数简称为月平滑黑子数。

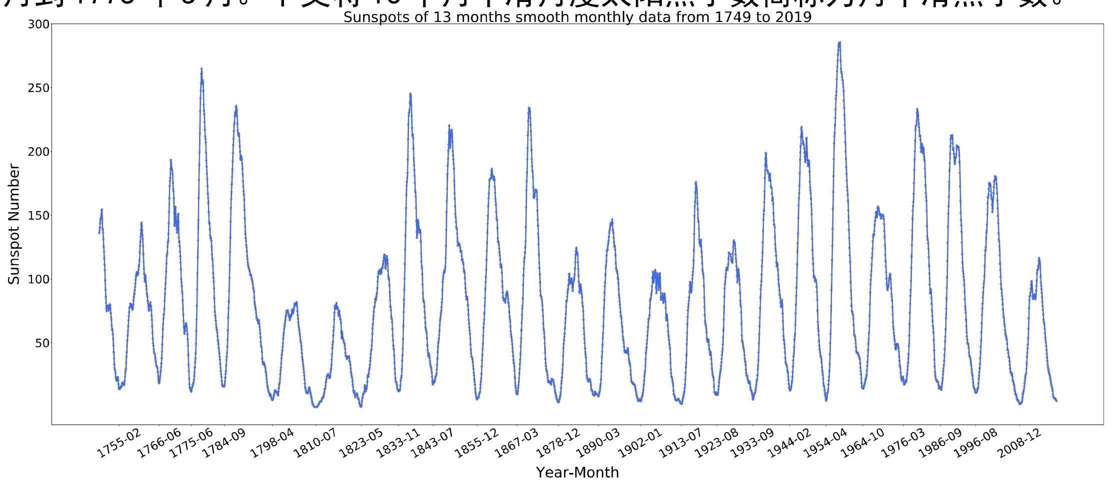


图 1　1749 年到 2019 年间的 13 个月平滑月度太阳黑子观测值

Fig. 1 13-month smooth monthly sunspot observations from 1749 to 2019

2. 方法

2.1 长短期记忆网络模型介绍

长短期记忆网络在神经网络快速发展的当今，被广泛应用到深度学习领域，并成为一种应用广泛的神经网络模型，与基础神经网络相比，长短期记忆网络处理时间序列的特性不仅在层与层之间建立了权连接，在相同层之间也建立了权链接。长短期记忆网络模型是一种循环神经网络模型，循环神经网络是一类以序列为基础，在序列的演进方向进行递归并将所有节点按链式链接的递归神经网络，这种链式链接的特性揭示了序列之间的密切关系。长短期记忆网络解决了循环神经网络在处理长序列数据时梯度消失的问题，所以，长短期记忆网络适用于处理时间序列中间隔和延迟相对较长的事件，例如语音识别、机器翻译以及时间序列的预测。

长短期记忆网络中的工作单元内部，如图 2，单元接收当前时刻的输入信息、上一时刻的隐藏状态和单元状态，并由 3 个门实现信息的持久化和抑制。它们分别是遗忘门、输入门和控制门。遗忘门通过激活函数决定了对前一时刻的隐藏状态“遗忘”多少信息。输入门包含两部分，（1）通过函数选择什么值要被

保留，（2）通过激活函数生成候选向量值，二者的乘积作为一部分状态量与遗忘门中生成的和上一时刻状态量的乘积之和作为当前的单元状态。最后，输出门通过生成候选向量，经过函数选择保留下来的信息，并将结果作为当前隐藏状态传输给下一个单元和上一层的同一时刻单元。

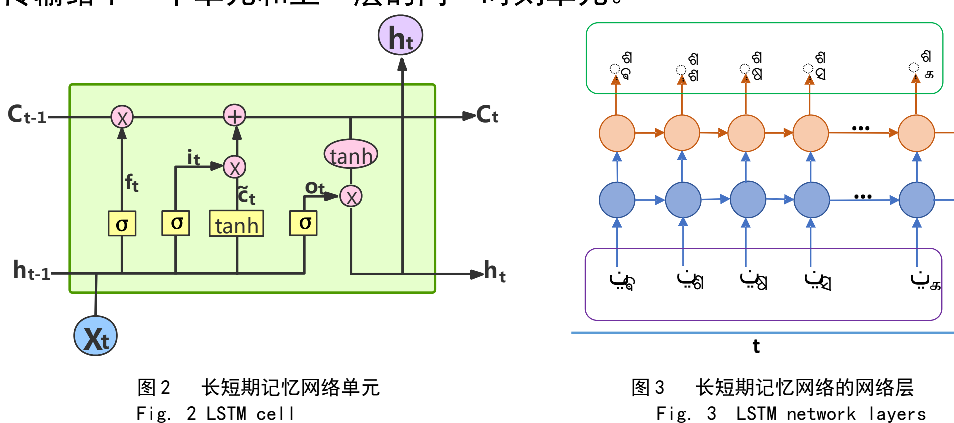


图 2　长短期记忆网络单元
Fig. 2 LSTM cell

图 3　长短期记忆网络的网络层
Fig. 3 LSTM network layers

长短期记忆网络的工作原理如图 3，是一个有个时间步长、两个网络层的结构，太阳黑子组成的时间序列, 第 1 层的 1 个工作单元在接收输入数据后与初始化的单元状态以及隐藏状态计算当前响应，并将响应传递给第 2 个单元以及上一层的第 1 个单元，第 1 层的第 2 个单元在下一时刻接收输入数据和上一个单元的状态量后, 又将结果传输给第 3 个单元和第 2 层的第 2 个单元，以此类推。第 2 层的单元接收第 1 层的输出作为输入，按照与第 1 层相同的方式传递计算的结果，并输出每个单元的隐藏状态作为每个时间步长的输出结果。在长短期记忆网络中，时间序列按照时间步长被分为多个有序数据段，每个数据段传入的工作单元中都以不同的权重计算输出结果。时间步长为 1 的预测方法称为单步长预测方法，大于等于 1 的称为多步长预测方法。

由此可知，当设时间步长为 1 的时候，只有一个 cell 提供权重和输出输入的响应，所有数据通过一个 cell 建立时间序列数据的信息持久性。而在多个时间步长的网络中，多个 cell 为时间序列提供不同的权重，并且不仅在 cell 自身建立信息的持久性，在链式链接的 cell 之间也为时间序列建立信息的持久化，并增加了时间序列数据在时间上的相关性。所以，从一定层度上说，多时间步长的网络模型能为时间序列在响应过程中提供更丰富的权链接和多重的信息相关性，基于此提出了多步长的预测方法。

表 1 时间步长策略

Table 1 Time step strategy

Multiple step & lag methods															
Step	1	2	3	4	5	6	8	10	12	15	20	24	30	40	60	120
Lag	120	60	40	30	24	20	15	12	10	8	6	5	4	3	2	1

2.2 多步长设计

为了验证多步长方法的效果并取得最优步长，提出了多组多时间步长的策略，如表 1。为了能够充分利用数据，所有的步长设置为 120 的约数，每组步长与滞后数之间都满足跨越 120 个数据的间隔，这与本文的目的相吻合，即预测 120 个月后的月平滑黑子数。时间序列按照时间步长分为多个有序数据段，滞后数表示每个数据段起始值之间所相隔的数据个数。步长为 1 滞后数为 120 的为单步长预测，其余 15 组均为多步长的设计方案。

2.3 多变量构建

在时间序列的预测范畴里，只用太阳黑子数据作为输入数据预测黑子数的方法属于单变量预测。除了上述的多步长预测太阳黑子周期，这里提出了一种加入太阳黑子周期长度构建多变量预测的方法，按照第 1 节提到的太阳黑子周期记录把每一个太阳黑子周期中的原黑子观测数据以他们在周期中的位置记录下来，作为除黑子原数据以外的另一个输入数据，把输入数据的维度增加到两维，这就是构建多变量的预测方法。具体方法举例：在一个黑子观测值个数为 135 的太阳周期中，将黑子在该周期中的观测值作为其中一个输入数据，并构建一个以 0 开头到 134 的序列，序列值与观测值一一对应。这不仅记录了观测值在周期中的相对位置，也记录了该周期的长度。月黑子平滑数在第 1 个最小周期开始之前有一段不完整的数据，用黑子周期的平均长 132 作为该不完整周期的长度，从 66 开始记录至 132 作为该 67 个观测值的序列。

代表太阳黑子的观测值，代表该观测时刻的序列值，代表观测数据的个数，多变量特征表示如下：

2.4 子序列采样

在整个太阳黑子的时间序列中，为了保留观察值对时间的依赖性，采用滑动窗口的方式选择连续的子样本，从而创建多个连续子序列分别构建模型，并在已有的观察值上评价开发策略。测试采用交叉验证的方法验证模型的预测效果，使得模型有更高的可靠性。

整个数据集分为两种长度的子序列，分别称为分片 11 和分片 6，分片 11 有 11 个子序列，分片 6 有 6 个子序列。每个子序列的数据被分为训练集和测试集，其中训练集用来训练长短期记忆网络模型，测试集用来评价长短期记忆网络预测模型。在深度学习方法中，为了使模型更健壮并且“学习”到更多的特征，在建模过程中使用的训练集数据量往往要比测试集庞大。当数据集的样本或观测值充足时，在保障测试集的数量能够有效验证模型并且满足实验目的的同时，尽可能地扩充训练集。本文用测试集数据量的 5 倍和 10 倍分别作为分片 11 和分片 6 的训练集。此外，训练集的数据量要大于等于 240，因为训练集会被划分为两部分：输入序列和目标值序列，二者起始值之间相差 120 个数据，并且无论输入序列的长度为多少，目标值序列的最小长度为 120。本文输入序列和目标值序列取相同长度，所以，当要预测未来 120 个月的太阳黑子数时，测试集的长度只能为 120。

分片 11 和分片 6 的训练分别包含 600 和 1200 个观测值，以及相同长度的测试数据 120，每个子序列之间的起始数据分别相隔 250 和 380 个观测值，采样参数如表 2。该采样方式虽然有两种长度的子序列，但两种分片跨越的数据总长度都是 3220 个，序列最后的 17 个数据没有被采用。在最后的预测工作中，训练数据由相同采样训练集的时间最邻近数据重新构建。

表 2 分片 11 和分片 6 的采样参数

Table 2　Sampling parameters of 11-slice and 6-slice

Slice number	Train length	Test length	Skip-span length	Remaining data
11	600	120	250	17
6	1200	120	380	17

2.5 验证方法

在已有的观测值上做“预测”的目的是为了给所设计模型的预测结果提供一个可靠的依据，所以，在月平滑黑子数据集中采样后的每一个子序列采用相同的模型以及模型中相同的超参数开发预测模型，将每一个子序列预测结果与

该时刻的观测值的均方根误差作为在此子序列上的预测效果。最终，对于一个完整的数据集，将所有子序列的平均均方根误差作为在该数据集上评判预测模型的效果。也就是说，对于11个子序列的采样方式，有11个均方根误差，对它们求均值，该均值就是在该种开发策略上的预测效果。

3. 实验结果和分析

用PyTorch框架作为开发模型，长短期记忆网络的设计上使用两个网络层，50个隐藏单元，并用Adam作为优化器，采用MAE作为损失函数。并用Visdom工具实时检测损失函数的收敛情况，模型一共训练100个周期。

结合第2.4节中的采样方式，对比了单变量和多变量所有步长方案，即表1中的预测效果，分别用所有子序列均方根误差的平均值（AVG）和所有子序列均方根误差的标准差（STD）来说明。

3.1 结果

3.1.1 单变量预测

首先对比了单变量在不同步长策略上的预测效果，其中包括单步预测和15种多步长的多步预测，表3分别展示了分片11和分片6在单步长预测和最优多步长上的平均均方根误差。分片11中步长为8、滞后数为15的多步长实验取得了最低的均方根误差平均值为43.1，相比单步长预测的45.0降低1.9，标准差从20.2降低到19.1。分片6中在步长为5、滞后数为24的多步长取得最低平均均方根误差值为50.4，相比单步长预测的53.5降低3.1，标准差反而从16.6增大到18.1。

表3 分片11和分片6的单变量在单步长和最优多步长的结果

Table 3 Univariate results of 11-slice and 6-slice in single-step and optimal multi-step

Methods			Subsequence											AVG	STD
			1	2	3	4	5	6	7	8	9	10	11		
11-slice	Univariate	Single-step	47.9	35.9	66.9	54.7	27.8	22.0	29.8	54.8	29.1	33.3	92.6	45.0	20.2
		Multi-step	49.9	31.6	68.0	35.2	16.6	31.8	33.7	46.2	36.9	34.8	89.1	**43.1**	**19.1**
6-slice	Univariate	Single-step	50.7	28.3	43.5	53.8	62.4	82.4	\	\	\	\	\	53.5	16.6
		Multi-step	48.4	24.6	37.6	50.1	58.6	83.2	\	\	\	\	\	**50.4**	18.1

3.1.2 多变量预测

多变量的预测同样对比了在单步长和多步长方法上的效果，表4分别展示了分片11和分片6在单步长预测和最优多步长上的平均均方根误差。分片11在步长为6、滞后数为20时的多步长取得最小均方根误差为42.8，相比单步预测的45.4降低2.6，标准差为15.6，比单步预测的23.6减小8.0。分片6在步长为12、滞后数为10时取得最小均方根误差为51.6，比单步预测的55.1降了3.5，标准差从20.1增大到23.4。

最后分别用分片11和分片6的最优策略，即多变量的6个步长和单变量的5个步长，预测了未来10年的太阳黑子数，结果见图4。在分片11中第25个太阳周期的最大振幅为144.9，出现在2022年12月。分片6的第25个太阳周期最大振幅为180.4，出现在2024年5月。

表4 分片11和分片6的多变量在单步长和最优多步长的结果

Table 4 Multivariate results for 11-slice and 6-slice at a single-step and optimal multi-step results

Methods			Subsequence											AVG	STD
			1	2	3	4	5	6	7	8	9	10	11		
11-slice	Multivariate	Single-step	47.1	26.6	94.7	57.5	15.4	17.1	42.0	52.0	29.9	38.2	79.2	45.4	23.6
		Multi-step	54.5	26.2	51.5	26.2	29.4	29.3	49.6	40.6	45.1	37.7	81.0	**42.8**	**15.6**
6-slice	Multivariate	Single-step	58.2	30.2	27.4	71.4	61.1	82.2	\	\	\	\	\	55.1	20.1
		Multi-step	63.0	20.5	28.5	46.6	60.4	90.9	\	\	\	\	\	**51.6**	23.4

图 4 分片 11(左)和分片 6(右)预测未来 10 年的太阳黑子数

Fig.4 11-slice (left) and 6-slice (right) predict the number of sunspots in the next 10 years

3.2 分析

从实验结果可以看出，就预测方法来说，无论是分片 11 还是分片 6，多步长预测结果都比单步长的好。在单变量和多变量的比较中，分片 11 的多步长平均均方根误差从 43.1 下降到 42.8，降低了 0.3，标准差降低了 3.5。对于分片 6，虽然多步长在多变量反而增加了 1.2，但是从结果可以看出，在第 2、3 和 4 子序列上结果都有明显的提升，但在第 1 和第 6 个子序列上误差过大使得最终的平均均方根误差大于单变量的多步长预测结果。综上所述，分片 11 在多变量的最优多步长，即步长为 6 时取得了最好的效果，不仅得到了更小的平均均方根误差，还有最小的均方根误差标准差。所以认为分片 11 的结果更为可靠，即未来 10 年的太阳黑子最大振幅为 144.9，出现在 2022 年 12 月。

4. 讨论

4.1 多变量的多步长预测效果

在上述结果中，分片 11 比分片 6 效果更突出，为此以分片 11 为例详细对比了每个子序列在单变量单步长、最优多步长和多变量最优多步长 3 种方法的均方根误差，如表 5。从每个子序列的均方根误差可以看出，在第 2、4、5、8、11 子序列中单变量的多步长预测比单步长预测误差小，多变量的多步长预测在第 2、3、4、8、11 子序列上误差均小于单变量的单步长和最优多步长预测结果。平均均方根误差在多变量的多步长预测中下降到了 42.8，相比单变量的单步长预测下降了 2.2，标准差则从 20.2 下降到 15.6，降低了 4.6。

表 5　分片 11 在单变量的单步长、最优多步长和多变量的最优多步长上的结果

Table 5 The results of 11-slice on univariate single-step, univariate optimal multi-step, and multi-variable optimal multi-step

Methods	Subsequence 1	2	3	4	5	6	7	8	9	10	11	AVG	STD
Univariate & Single-step	47.9	35.9	66.9	54.7	27.8	22.0	29.8	54.8	29.1	33.3	92.6	45.0	20.2
Univariate & Multi-step	49.9	**31.6**	68.0	**35.2**	**16.6**	31.8	33.7	**46.2**	36.9	34.8	**89.1**	**43.1**	**19.1**
Multivariate & Multi-step	54.5	**26.2**	**51.5**	**26.2**	29.4	29.3	49.6	**40.6**	45.1	37.7	**81.0**	**42.8**	**15.6**

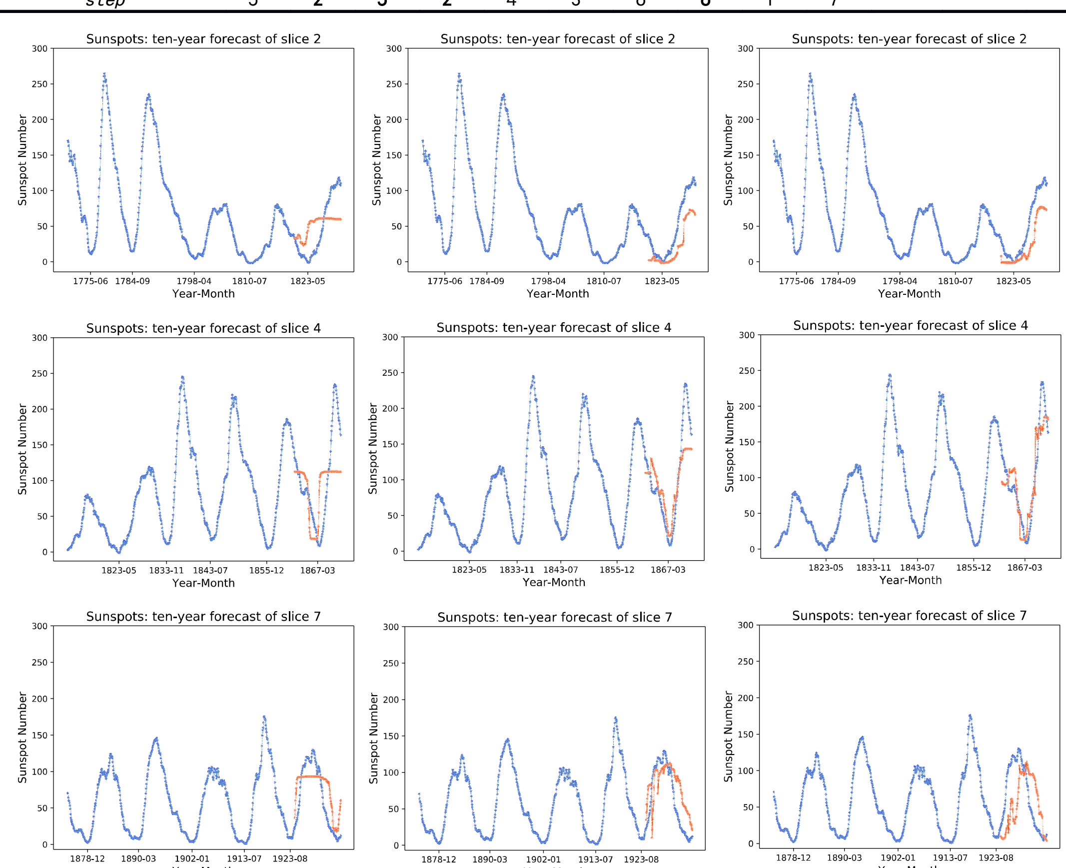


图 5 第 2、4 和 7 个分片在单变量的单步长、单变量的多步长和多变量的多步长预测结果

Fig.5 The results of slice 2, 4 and 7 prediction in univariate single step, univariate multi-step and multivariable multi-step respectively

本文详细列举了第 2、4 和 7 个子序列在 3 种方法中的预测结果。如图 5，每一行分别代表 3 个不同的子序列，每一列分别代表单变量的单步、单变量的多步和多变量的多步三 3 预测方法。在单变量的单步预测中，预测的黑子变化过于平滑，在预测的极大值处很容易出现连续的水平预测值。在单变量的多步预测中，极值处的水平变化有所改善却又波动过大。相比其他两种方法，多变量的多步长预测不仅没有在极大值处于水平变化，也没有周期内的大幅度波动，

反而较为准确地预测了周期的起始和结束时间(第 7 个子序列)以及在该时刻的极小值，同样也较准确预测了最大振幅的数值和出现的时间。

第 7 个子序列的均方根误差在 3 种方法中逐渐增大，但是通过对预测结果的最大振幅、周期起始点和结束点在数值和时间上的分析，断定在最大的均方根误差也就是多变量的多步长预测中有更好的预测效果，造成这种现象的原因应该是两个极值之间的预测过程中出现的波动以及相位的偏移，子序列 5、6 与 7 情况类似。第 1、9、11 个子序列则由于所预测的周期峰值突变，造成了较大的均方根误差。

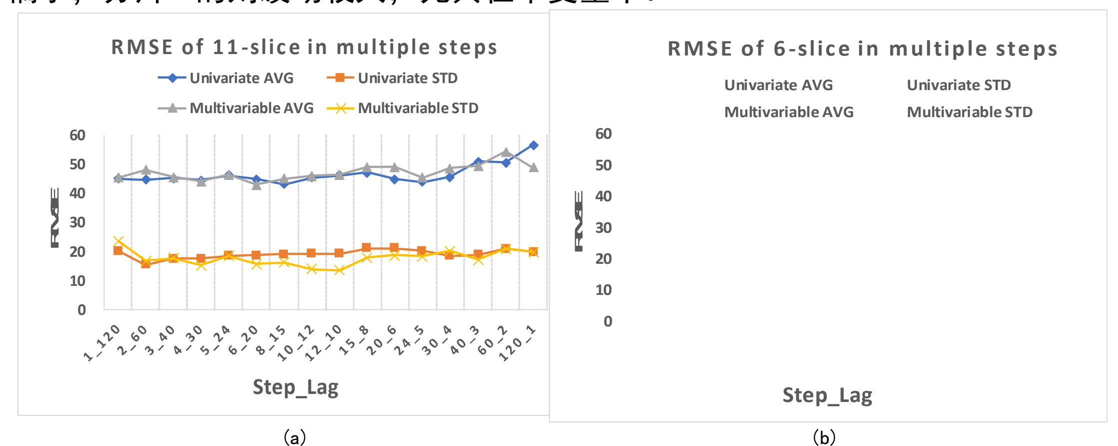
4.2 数据集与采样方法对预测结果的影响

文[16]，文[20-21]均采用了长短期记忆网络模型预测未来的太阳黑子数，其中文[16]结合 SILSO 的太阳黑子月度平均数据及单变量的单步长方法预测未来 10 年的太阳黑子数，并把数据分为 6 和 12 个子序列训练集长度分别为 1200 和 840，测试集分别为 600 和 240，平均均方根误差分别为 35.9，36.9。文[21]用来自 R 数据集的太阳黑子月度数据及单变量的 12 个步长方法预测未来 12 个月的太阳黑子，相似地把数据分为 6 个子序列，训练集和测试集长度分别为 1200 和 600，平均均方根误差为 26.9。文[20]使用与文[16]相同的数据集及单变量的单步长方法预测未来 10 年的太阳黑子数，将数据分为 11 个子序列，训练数据和测试数据分别为 600 和 120 个，得到平均均方根误差为 34.4。在文[16]与文[21]训练数据和测试数据为 1200 和 600 的结果比较中，平均均方根误差相差 9.0。本文的分片 11 与文[20]有相同长度的训练和测试数据，取得的平均均方根误差为 45.0，与文[20]相差 10.6。

数据集本身的差异以及采样的训练、测试数据长度不一致，导致结果有较大的差异，其中采样方法不一致使均方根误差在一定程度上没有办法类比。虽然 R 数据集有更好的预测效果，但是因为数据只记录到 2013 年，致使无法采用。SILSO 的太阳黑子数据与 R 数据集的太阳黑子数据相比更离散，黑子的波动幅度更大，致使预测结果均方根误差较差。

4.3 分片 11 和分片 6 在不同步长中的预测结果

本文对单变量和多变量在表 1 提出的所有时间步长策略进行了实验，如图 6。结果证明，随着步长的增加，平均均方根误差呈现下降趋势，这种下降趋势在分片 6 中更为明显，之后随着步长的大幅度增加，误差越来越大，并且分片 11 的平均均方根误差总体比分片 6 的偏小。在所有分片的均方根误差标准差中，分片 11 随着步长的增多变化相对平稳，并且多变量的实验结果比单变量的总体偏小，分片 6 的则波动较大，尤其在单变量中。



(a) (b)

图 6 分片 11(a)和分片 6(b)在多种步长策略的预测结果

Fig.6 Prediction results of 11-slice (left) and 6-slice (right) in multiple step strategies

由此可见，从总体的角度来看分片 11 确实是比分片 6 更好的采样方式，并且在分片 11 中多变量的预测结果确实更好。

4.4 在第 21、22、23 和 24 太阳周期上的预测结果

为了给预测结果提供可靠的论证，用分片 11 和分片 6 的最优策略预测了第 21、22、23 和 24 太阳周期的月平滑黑子数，图 7 直观地展示了预测效果，并将结果详细地记录在表 6 中，并用“振幅”表示最大振幅。

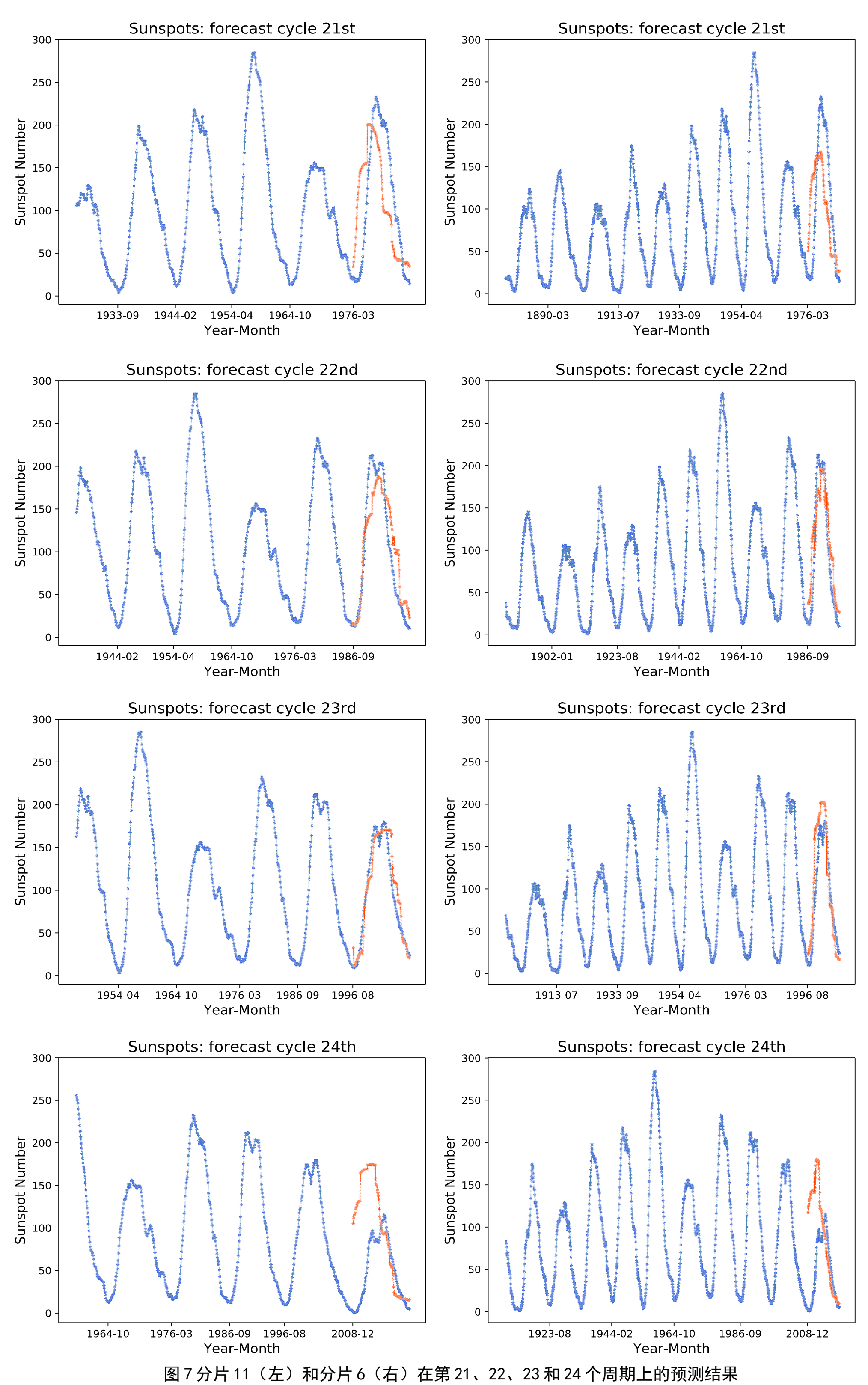


图 7 分片 11（左）和分片 6（右）在第 21、22、23 和 24 个周期上的预测结果

Fig.7 Prediction results of 11-slice (left) and 6-slice (right) in cycles 21, 22, 23 and 24

在第 21、22、23 和 24 周期的预测结果中，分片 11 的最大振幅与真实记录相比分别降低了 12.7%、11.5%、6.7%和增大了 50%，最大振幅的出现时间与真实记录相比分别提前了 1 年、推迟了 1 年 2 个月、推迟了 1 年 1 个月、提前了 2

年1个月；在分片6中最大振幅与真实记录相比分别降低了23.4%、7.5%、增大了12.4%和54.9%，最大振幅的出现时间与真实记录相比分别推迟了2个月、推迟了9个月、提前了11个月和提前了2年9个月。

表6 分片11和分片6在第21、22、23和24太阳周期上的预测结果

Table 6 Prediction results of 11-slice and 6-slice in the 21st, 22nd, 23rd and 24th solar cycles

Cycles	21st	22nd	23rd	24th
True value (Amplitude / Time)	232.9/1979.12	212.5/1989.11	180.3/2001.11	116.4/2014.4
11-slice (Amplitude / Time /RMSE)	203.4/1978.12/59.1	187.9/1991.1/33.4	168.2/2002.12/21.4	174.6/2012.3/72.0
6-slice (Amplitude / Time /RMSE)	178.3/1980.2/59.5	196.5/1990.8/32.5	202.7/2000.12/29.3	180.4/2011.7/68.1

由以上数据分析可以看出，分片11的预测结果在最大振幅上的波动除了第22周期都要比分片6小，并且4个周期的平均均方根误差为46.5比分片6的47.4略小。而分片6在最大振幅出现的时间都表现出更高的准确度，除了第24周，出现这种现象的原因：在构建模型的训练数据中分片6包含的完整周期比分片11多，所以“学习”到的周期长度特征更多，对此能够将时间跨度预测的更准确。在第24太阳周期预测中，振幅和周期起始时间偏差都过大，振幅偏差过大是因为与之前的周期相比第24周期的振幅骤然下降，从而导致了偏差较大的振幅；而周期起始时间偏差过大则是因为第23周期的历时相比之前的周期来说要更久，有12年4个月，从而直接影响对下一个周期起始时间的预测。

总的来说，无论是分片11还是分片6，在各自的最优策略上取得的预测效果都不错，分片11在振幅的预测上有更大的优势，分片6在预测周期的历时上更为准确，这为本文预测未来10年的太阳黑子结果提供了有力的依据。

4.5 本文方法在太阳黑子预测上的局限

从上面的讨论可以看出，子序列的采样方式直接决定了数据特征对长短期记忆网络模型的构建，对预测结果有非常大的影响，这使子序列采样方法成为对比不同预测方法中的一项苛刻条件。在第4.4节的分析中，第24太阳周期振幅突变造成预测结果偏差较大，使得结果的准确性建立在下一周期没有发生突变的情况下。要为预测结果提供可靠的论证，只能将之前周期的“预测”结果作为参考，这就要求数据集的观测量不能太小，并且要合理设计子序列的长度，保证能够在训练出预测模型的同时，还能为该方法的效果提供合理的论证。

图5中的第7个分片预测的恰好是一个完整的周期，可以看出该周期有双峰结构，在图7中的第22、23和24太阳黑子周期也都有明显的双峰结构，但在预测的结果中它们多以最大振幅处水平变化或者是单峰振幅并出现多个折点的形式表现，并没有明显的双峰结构。最大振幅处的水平变化说明训练好的模型确实有保留双峰的特性，但两个峰值之间的变化不明显，致使模型在训练过程中没有“学习”到双峰的细微特征而出现水平变化。

5. 总结

本文使用深度学习的经典模型长短期记忆网络开发了预测月平滑黑子数的模型，并提出了一种构建多变量的方法。通过比较两种采样方法分片11和分片6在多变量和单变量的16组步长上的平均均方根误差，在分片11的多变量并且步长为6时取得最小的平均均方根误差为42.8。并在详细的结果分析中，通过3个点的准确度证明多变量的多步长预测确实有更好的效果，3个点分别是周期

的起始位置、周期的结束位置和最大振幅的位置。分片 11 和分片 6 在多变量和单变量的 16 组步长预测结果分析中证明分片 11 的采样策略比分片 6 有更好的稳定性和更低的均方根误差。在对第 21、22、23 和 24 周期的“预测”中，分片 11 表现出对预测最大振幅的相对优势，而分片 6 则在最大振幅出现的时间上表现得更准确，这些结果为预测未来 10 年的太阳黑子数提供了可靠的依据。换而言之，在分析的过程中也能发现该预测方法的局限，例如在预测第 24 太阳周期中振幅突变、上一周期历时突变导致预测结果偏差较大；以及模型不够完善在预测的结果中没有体现出周期的双峰特性，这或许是该领域接下来的研究重点。

最终，以分片 11 的多变量在步长为 6 处的模型预测了未来 10 年的太阳黑子活动，即第 25 个太阳活动周期，最大振幅为 144.9，出现在 2022 年 12 月，比第 24 周期活跃，持续时间至少为 10 年。

参考文献:

[1] 苏同卫, 李可军. 太阳活动对中国中纬度地区 8 级大地震的可能触发[J]. 天文研究与技术, 2007, 4(2):195-198.

[2] 沈玫, 吕达仁. 太阳活动与天气和气候变化的关系的评述[J]. 天文研究与技术, 1989(S1):242-254.

[3] DU Z L, LI R, WANG H N.The predictive power of ohl’s precusor method[J]. The Astronomical Journal, 2009, 138(6):1998-2001

[4] SCHATTEN K H, SOFIA S. Forecast of an exceptionally large even-numbered solar cycle[J]. Geophysical Research Letters, 2013, 14(6):632-635.

[5] SARP V, KILCIK A, YURCHYSHYN V, et al. Prediction of solar cycle 25: a non-linear approach[J]. Monthly Notices of the Royal Astronomical Society, 2018, 481(3): 2981-2985.

[6] KILCIK A, ANDERSON C N K, ROZELOT J P, et al. Nonlinear prediction of solar cycle 24[J]. The Astrophysical Journal, 2008, 693(2):1173-1177.

[7] AGUIRRE L A, LETELLIER C, MAQUET J. Forecasting the time series of sunspot numbers[J]. Solar Physics, 2008, 249(1): 103-120.

[8] MARIS G, ONCICA A. Solar cycle 24 forecasts[J]. Sun & Geosphere, 2006, 1:8-11.

[9] KITIASHVILI I, KOSOVICHEV A G. Application of data assimilation method for predicting solar cycles[J]. The Astrophysical Journal, 2008, 688(1):L49-L52.

[10] KITIASHVILI I N, KOSOVICHEV A G. Nonlinear dynamical modeling of solar cycles using dynamo formulation with turbulent magnetic helicity[J]. Geophysical & Astrophysical Fluid Dynamics, 2009, 103(1):53-68.

[11] PETROVAY, KRISTÓF.Solar cycle prediction[J]. Living Reviews in Solar Physics, 2010, 7(1):1-59.

[12] SARP V, KILCIK A, YURCHYSHYN V, et al. Prediction of solar cycle 25: a non-linear approach[J]. Monthly Notices of the Royal Astronomical Society, 2018, 481(3): 2981-2985.

[13] BHOWMIK P, NANDY D. Prediction of the strength and timing of sunspot cycle 25 reveal decadal-scale space environmental conditions[J]. Nature Communications, 2018, 9(1).

[14] SELLO S. Solar cycle activity: an early prediction for cycle #25[J]. 2019.

[15] PESNELL W D, SCHATTEN K H. An early prediction of the amplitude of Solar Cycle 25[J]. Solar Physics, 2018, 293(7): 112.

[16] PALA Z, ATICI R. Forecasting sunspot time series using deep learning methods[J]. Solar Physics, 2019, 294(5).

[17] ATTIA A F, ISMAIL H A, BASURAH H M. A neuro-fuzzy modeling for prediction of solar cycles 24 and 25[J]. Astrophysics and Space Science, 2013, 344(1):5-11.

[18] DANI T, SULISTIANI S. Prediction of maximum amplitude of solar cycle 25 using machine learning[C]//Journal of Physics: Conference Series. IOP Publishing, 2019, 1231(1): 012022.

[19] HOCHREITER, SEPP, SCHMIDHUBER, JÜRGEN. Long short-term memory[J]. Neural Computation, 9(8):1735-1780.

[20] https://www.r-bloggers.com/time-series-deep-learning-forecasting-sunspots-with-keras-stateful-lstm-in-r/

[21] DANCHO & KEYDANA. TensorFlow for R: Predicting Sunspot Frequency with keras. Retrieved fromhttps://blogs.rstudio.com/tensorflow/posts/2018-06-25-sunspots-lstm/

Prediction and analysis of sunspot activity based on multivariable LSTM network

Liang Bo, Lin Yuqi, Dai Wei, Feng Song, Yang Yunfei
(1.Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China；
2.Yunnan Key Laboratory of computer application, Kunming 650500, China)

Abstract: By adding the length of sunspot cycle, the multivariable input data of LSTM is constructed to predict the change of sunspot number in the next ten years with multiple time steps. According to the length of training data, the data set is divided into two groups of time series segments, namely, 11-slice and 6-slice. The prediction effect of univariate and multivariable on single-time step and multi-time step is compared. Finally, the main results are as follows:(1) Compared with the method of 6-slice, the method of 11-slice has lower root mean squared error (RMSE). (2) The optimal multi-step prediction is better than the single-step. (3) The accuracy of the starting point, ending point, and the maximum amplitude position of the cycle on the legend proves that prediction effect of the multivariate multi-step method is better than that of single-step method.

Keywords: Sunspot; LSTM; Multivariate; Time step; RMSE